JOURNAL OF VIROLOGY, Dec. 1994, p. 7717–7727
0022-538X/94/$04.00+0

Vol. 68, No. 12

Structure and Function of the Murine Cytomegalovirus *sgg1* Gene: a Determinant of Viral Growth in Salivary Gland Acinar Cells

LAUREL A. LAGENAUR,[1] WILLIAM C. MANNING,[1]† JEFFREY VIEIRA,[1]‡
CHRISTINE L. MARTENS,[2]§ AND EDWARD S. MOCARSKI[1]*

Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305-5402,[1] and DNAX Research Institute, Palo Alto, California 94304-1104[2]

Received 10 January 1994/Accepted 31 August 1994

The salivary gland has long been recognized as an important target organ for cytomegalovirus replication in the infected host. A viral gene, denoted *sgg1*, plays an important role for replication in the salivary gland even though it is dispensable for growth in other organs or in cultured cells. The nucleotide sequence of this gene and of cDNA clones representing two spliced transcripts (1.5 and 1.8 kb in size) has been determined. The more abundant 1.5-kb transcript contains a 312-amino-acid (aa) open reading frame (ORF) and encodes the corresponding 37-kDa protein (Sgg1) when expressed in transfected COS-7 cells. The 1.8-kb transcript initiates upstream of the 1.5-kb transcript and contains a 108-aa ORF in addition to the 312-aa ORF. This longer cDNA also encodes the 37-kDa protein Sgg1, although at lower abundance than the 1.5-kb cDNA. Sgg1 localizes to the cytoplasm of COS-7 cells, which is consistent with the predicted structural characteristics of the 312-aa ORF: a type 1 integral membrane protein. During viral infection, expression of both *sgg1* transcripts is highest at early times (8 to 12 h) after infection; only the 1.5-kb transcript is present, at low levels, late in infection. A recombinant virus, RM868, carrying a *lacZ-gpt* insertion within *sgg1*, fails to express Sgg1 protein and exhibits reduced growth in the salivary gland. RM868 retains the capacity to disseminate in the infected mouse and to enter serous acinar cells, although it fails to replicate efficiently in this cell type. These results suggest that *sgg1* is critical for high levels of viral replication in the salivary gland.

Cytomegalovirus (CMV) is an important human pathogen in utero and in immunocompromised individuals. Human CMV causes only a mild febrile illness in immunocompetent individuals; however, infection of an immunocompromised host, such as a developing fetus, a transplant recipient, or an individual with AIDS, may result in severe, life-threatening illness (1, 17). Human CMV and related CMVs that infect other animal species (35) are all highly species specific, yet their genome organization and replication characteristics are similar (17, 32). These similarities have led to the study of animal viruses such as murine CMV as models of human CMV (18, 20, 33). A key feature of all herpesviruses is the ability to replicate persistently and establish latency in specific cell types in the host. For CMVs, epithelial cells in the salivary glands and kidneys and mononuclear cells in blood have been implicated as sites of replication and latency. Because the mouse is its natural host, viral determinants of tissue tropism can be identified and studied with murine CMV. Procedures for construction of recombinant viruses carrying specific mutations have fostered the identification of such determinants (23, 24, 44, 47).

CMVs have a well-established tropism for the salivary gland (8, 40, 41). Virus produced there and shed in saliva serves as an important source for transmission (36). The mouse has three major sets of salivary glands—the submaxillary, parotid, and sublingual—located in the subcutaneous tissues of the neck (10, 39). Submaxillary glands, the largest of the three and the principal site of murine CMV replication, are composed of three major epithelial cell types: serous acinar, mucous acinar, and ductal. Murine CMV replicates predominantly in serous acinar epithelial cells (14, 26, 31), in contrast to human CMV, which replicates in the ductal epithelium (17) of the salivary glands. Murine CMV produced in the serous acinar cells is more highly virulent than cell-culture-passaged virus. Just one passage of cell-culture-propagated virus through the mouse, however, restores virulence (19, 34). After any inoculation route and with either cell-culture-propagated or mouse-passaged virus, the highest titers of virus are consistently found in the submaxillary salivary gland.

Manning et al. (24) previously described a recombinant murine CMV, RQ401, carrying a mutation in a gene that was denoted *sgg1*. The *sgg1* gene is located adjacent to the *ie2* gene (29), within the *Hin*dIII J fragment. *sgg1* is dispensable for growth in cultured cells as well as for growth in the spleen and liver of an infected mouse; however, *sgg1* mutants replicate poorly in the salivary gland of the mouse, with peak titers reduced by 3 to 4 orders of magnitude. The spontaneous mutation originally mapped in *sgg1* mutant RM427 and its derivative RQ401 disrupted expression of two overlapping β transcripts: a major 1.5-kb species and a minor 1.8-kb species (24). The mutation was rescued in recombinant virus RQ427, restoring expression of the two *sgg1* transcripts and recovering the ability of the virus to replicate to high levels in the salivary gland. The mutation causing this phenotype was identified as a 323-bp deletion, which removed a predicted polyadenylation signal shared by both transcripts.

Here, we report the nucleotide sequence of the *sgg1* region. We examine the RNA expression kinetics and determine the start sites of transcription for both the 1.5- and 1.8-kb transcripts. We identify the principal gene product of *sgg1* and determine its subcellular localization. We describe an additional *sgg1*-deficient virus, RM868, and show the structure of this virus and its growth phenotype in the mouse. Finally, we

* Corresponding author. Phone: (415) 723-6435. Fax: (415) 723-9180. Electronic mail address: MK.ESM@Forsythe.Stanford.EDU.
† Present address: Chiron Corp., Emeryville, CA 94608.
‡ Present address: Department of Laboratory Medicine, University of Washington, Seattle, WA 98144.
§ Present address: Affymax Research Institute, Palo Alto, CA 94304.

TABLE 1. Plasmid constructs used in this study

Plasmid	Source fragment/plasmid	Cloning vector/plasmid
pON450	*Xho*I-*Sna*BI/pON446	*Sal*I-*Sma*I/pGEM2
pON451	*Stu*I-*Xho*I/pON446	*Sal*I-*Sma*I/pGEM2
pON452	*Stu*I-*Sna*BI/pON446	*Sma*I/pGEM2
pON457	*Cla*I-*Xho*I/pON446	*Cla*I-*Xho*I/Bluescript KS$^-$
pON478	*Eco*RI-*Stu*I/pON446	*Eco*RI-*Sma*I/pGEM3Zf$^+$
pON482	*Xho*I-*Sna*BI/pON446	*Sal*I-*Sma*I/pGEM3Zf$^+$
pON483	*Xho*I-*Stu*I/pON446	*Sal*I-*Sma*I/pGEM3Zf$^-$
pON487	*Cla*I-*Eco*RI/pON459	*Acc*I-*Eco*RI/pGEM3Zf$^-$
pON4000	*Cla*I-*Nae*I/pON446	*Acc*I-*Sma*I/pGEM3Zf$^+$
pON4001	*Nae*I-*Cla*I/pON446	*Sma*I-*Acc*I/pGEM3Zf$^+$
pON4003	*Cla*I-*Nae*I/pON446	*Acc*I-*Sma*I/pGEM3Zf$^+$
pON4005	*Eco*RI-*Stu*I/pON446	*Eco*RI-*Sma*I/pGEM3Zf$^-$
pON4010	*Sac*II-*Cla*I/pON459	*Sac*II-*Cla*I/Bluescript KS$^+$
pON4011	*Sac*II-*Cla*I/pON459	*Sac*II-*Cla*I/Bluescript SK$^+$
pON4012	*Stu*I-*Sac*II/pON459	*Sma*I-*Sac*II/Bluescript SK$^+$
pON4013	*Stu*I-*Sac*II/pON459	*Sma*I-*Sac*II/Bluescript KS$^+$

investigate the nature of the inhibition of growth in salivary glands exhibited by RM868 by using histochemical staining.

MATERIALS AND METHODS

Cells and viruses. Wild-type strains K181 (from M. Colin Jordan) and Smith (ATCC VR-194) and recombinant viruses RQ401, RQ427 (24), RM461 (44), and RM868 (47) were grown on NIH 3T3 fibroblast cells (ATCC CRL1658) in Dulbecco's modified Eagle's medium (GIBCO-Life Technologies, Bethesda, Md.) supplemented with 10% NuSerum (Collaborative Research, Waltham, Mass.) (42, 44). K181$^+$ is a plaque-purified virus derived from K181 (44). COS-7 cells (obtained from DNAX Research Institute, Palo Alto, Calif.) used for transient expression assays were grown under the same conditions.

Recombinant plasmids. Plasmid cloning was by standard methods (37). Vectors pGEM3Zf$^+$ and pGEM3Zf$^-$ (Promega, Madison, Wis.) and Bluescript SK$^+$ and Bluescript KS$^+$ (Stratagene, La Jolla, Calif.) were used for single-stranded DNA sequence determination, and pGEM2 (Promega) was also used for double-stranded sequence determination. All restriction enzymes were purchased from New England Biolabs (Beverly, Mass.) and used with buffer supplied by the manufacturer. The plasmids that were constructed for DNA sequence analysis are depicted in Table 1.

In addition to the plasmids shown in Table 1, we used the following plasmids as previously described (24). pON446 carries an *Eco*RI-*Hin*dIII fragment from K181 DNA, pON455 carries an *Msc*I-*Cla*I fragment from pON446 inserted into *Hinc*II-*Acc*I-digested pGEM2, pON484 and pON456 carry a *Cla*I-*Xho*I fragment of pON446 cloned into *Cla*I-*Xho*I of Bluescript KS$^+$ and KS$^-$, pON486 carries an *Msc*I-*Xho*I fragment from the pON446 fragment inserted into *Sma*I-*Sal*I-digested pGEM3Zf$^-$, pON477 and pON481 carry an *Msc*I-*Hin*dIII fragment from pON446 cloned into *Hinc*II-*Hin*dIII pGEM3Zf$^+$ and pGEMZf$^-$, pON4002 carries a *Nae*I-*Cla*I fragment from pON446 cloned into *Sma*I-*Acc*I-digested pGEM3Zf$^-$, and pON4004 carries a *Hin*dIII-*Cla*I fragment from pON446 cloned into *Hin*dIII-*Acc*I-digested pGEM3Zf$^+$. pON459, an *Eco*RI-*Bam*HI genomic fragment from pON449 (*Hin*dIII J), was inserted into pUC9 (22). pON868 carries an insertion of a *lacZ-gpt* cassette in a *Sau*3A site located within a 740-bp *Cla*I fragment in the *sgg1* gene (47).

Nucleotide sequence analysis. Single-stranded DNA was prepared according to a Promega protocol (35a), except that RNase (Boehringer Mannheim, Indianapolis, Ind.) was added to supernatant phage after harvesting and samples were incubated at 37°C for 30 min. T7, T3, and SP6 promoter primers (Promega) or synthesized 15-bp oligonucleotides were annealed to single-stranded or denatured double-stranded template. All synthesized oligonucleotides were prepared on a model 391 DNA/RNA synthesizer (Applied Biosystems, Foster City, Calif.). The following oligonucleotides were prepared as primers for sequence determination: 4011-3′, 5′GGTTGA TATCTACTC3′; 4010-5′, 5′CAGACGAGCAGCTCCC3′; 4011-5′, 5′GGGAGCTGCTCGTCTG3′; 4013, 5′CACCACT CGATGTAG3′; 4012, 5′CTACATCGAGTGGTG3′; 4013-2, 5′GATGATCAGATCGATC3′; 4012-2, 5′GTTGGTCGGAG CCGA3; 4012-3, 5′GACCGCGACCGTTCG3′; and 4013-3, 5′CCACCGACCATCTAC3′. Sequenase version 2.0 (U.S. Biochemical Corp., Cleveland, Ohio) was used according to the manufacturer's protocol. Samples were separated on 8% polyacrylamide–urea gels as described previously (37). [^{35}S] dATP-labeled DNA was detected on radiographic X-Omat film (Kodak, Rochester, N.Y.). Determination of the DNA sequence of pON868 was performed with the *fmol* DNA sequencing system (Promega) with *sgg*580f primer (5′ACGGT GGAGACCGTCAATCA3′) according to the manufacturer's protocol. DNA sequence analysis was performed with the Wisconsin Genetics Computer Group package of programs run on a Sun Sparc workstation.

RNA preparation and RNA blot analysis. Whole-cell RNA was harvested from NIH 3T3 cells infected with murine CMV K181$^+$ or K181 virus by the guanidinium isothiocyanate procedure (5, 11). Five to 10 μg of RNA was separated on a 1% agarose–6% formaldehyde gel, transferred to either nitrocellulose (Schleicher & Schuell, Keene, N.H.) or GeneScreen Plus (NEN Research, Boston, Mass.), and hybridized as described previously (11, 16).

cDNA library construction. Whole-cell RNA was harvested at 8 h postinfection (hpi) as described above, and 50 μg of poly(A) RNA was isolated with oligo(dT)-cellulose (Collaborative Research, Bedford, Mass.). A cDNA library was prepared with a cDNA synthesis kit (Pharmacia LKB, Piscataway, N.J.), according to the manufacturer's protocol. After synthesis, cDNAs were ligated to *Bst*XI adapters (Invitrogen, San Diego, Calif.) and size fractioned over a Sephacryl S-400 column (Pharmacia) to exclude free adapters. cDNAs were then cloned into an expression vector, pME18S (25), derived from pCD-SRα (45). The resulting library was transformed into *Escherichia coli* MC161/WM1100 (Bio-Rad Laboratories, Richmond, Calif.). Bacterial stock titers were determined, and bacteria were plated onto large petri dishes (150-mm diameter) at a concentration of 10^3 colonies per plate. After growth, the colonies were transferred to nitrocellulose and bacteria were lysed in 10% sodium dodecyl sulfate (SDS). DNA was denatured in 0.5 M NaOH–1.5 M NaCl solution and neutralized in 0.5 M Tris-Cl (pH 8.0)–1.5 M NaCl buffer. DNA was then UV cross-linked to the nitrocellulose by UV Stratalinker 1800 (Stratagene; 0.12 J/cm^2) for 2 min and hybridized at 65°C in hybridization solution (6× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 1× Denhardt's solution, and 100 mg of salmon sperm DNA per ml) overnight. The cDNA library was probed with [α-^{32}P]dCTP-labeled random hexamer-primed (Pharmacia) fragments as described previously (37). The library was screened with a 592-bp *Cla*I fragment from pON446 that detects both the 1.5-kb cDNA and the 1.8-kb cDNA and a 383-bp *Sac*II-*Cla*I fragment from pON459 that was specific to the 1.8-kb cDNA. The resulting cDNAs ob-

tained from library screening were named pON4014 (1.5-kb cDNA), pON4019 (1.5-kb cDNA in reverse orientation), and pON4015 (1.8-kb cDNA).

Animals and antibodies. Five-week-old female CD-1 mice were purchased from Charles River (Wilmington, Mass.) and inoculated intraperitoneally (i.p.) with 10^6 PFU of tissue-culture-propagated murine CMV (Smith strain). Three-week-old male BALB/c.ByJ mice were purchased from Jackson Laboratories (Bar Harbor, Maine) and were inoculated i.p. with 10^6 PFU of tissue-culture-propagated murine CMV ($K181^+$, Smith, or RQ401). Mice were inoculated i.p. twice at 1-month intervals with 10^6 PFU. Blood was collected from the tail vein, and sera were tested individually for reactivity to the Sgg1 protein by immunoprecipitation.

Primer extension and RNase/S1 protection. A 35-bp oligonucleotide primer was synthesized and denoted 970-936 (5'CA GGTCAGCACGATGGTCTGAGCGGACACGGTCTC3'). This primer was [γ-^{32}P]ATP end labeled with polynucleotide kinase (New England Biolabs) and hybridized with 10 μg of whole-cell RNA (from $K181^+$-infected NIH 3T3 cells) for 10 min at 25°C. Reverse transcription of this RNA was performed with deoxynucleoside triphosphates (Pharmacia) at a concentration of 1 mM (each), 1 U of RNasin (Promega), 200 U of SuperScript RNase H reverse transcriptase (GIBCO BRL, Gaithersburg, Md.), and 1 ml of 0.1 M dithiothreitol (GIBCO) at 44°C for 90 min, followed by treatment with RNase (Boehringer Mannheim) for 30 min at 37°C. Primer extension samples were electrophoretically separated on an 8% polyacrylamide–urea gel.

A 228-bp probe was generated for S1 nuclease analysis by digesting pON459 with *Eco*RI and dephosphorylating and labeling the 5' end with [γ-^{32}P]ATP by using polynucleotide kinase. The DNA was then digested with *Cla*I, and the resulting fragment was purified, denatured, and added to 50 μg of whole-cell RNA harvested from $K181^+$-infected or mock-infected cells. The mixture was heated for 15 min, hybridized overnight at 58°C, and digested with 300 U of S1 nuclease (GIBCO BRL). The products were extracted with phenol-chloroform, ethanol precipitated, and resolved by electrophoresis on an 8% polyacrylamide gel containing 7 M urea and visualized by autoradiography.

Transient transfection assay. Three micrograms of plasmid DNA was transfected into 3×10^5 COS-7 cells by a calcium phosphate method as previously described (4), except that the incubation of the cells with the calcium phosphate-DNA mixture was in a 5% CO_2 incubator. Cells were allowed to recover for 24 h, and immunoprecipitation was performed.

Immunofluorescence. COS-7 cells were seeded into chamber slides (Labtek, NUNC, Inc., Naperville, Ill.) and transfected with either pON4014 or pON4015, the 1.5- or 1.8-kb cDNA. At 32 h posttransfection, the cells were washed with cold phosphate-buffered saline (PBS) and fixed with acetone-methanol in a 50:50 ratio. Slides were preincubated for 10 min with PBS containing 1% fetal bovine serum at 25°C. Mouse anti-$K181^+$ or anti-RQ401 antiserum was added to fixed cells at various dilutions (1:50 to 1:250) in PBS–1% fetal bovine serum and incubated for 1 h at 25°C. Slides were washed with the PBS–1% fetal bovine serum mixture, and goat anti-mouse immunoglobulin G fluorescein isothiocyanate conjugate (Cappel, West Chester, Pa.) was added at a dilution of 1:80 in PBS–1% fetal bovine serum. Slides were then incubated for 1 h at 25°C, in the dark, in a humidifying chamber. Samples were examined with a Zeiss epifluorescence microscope.

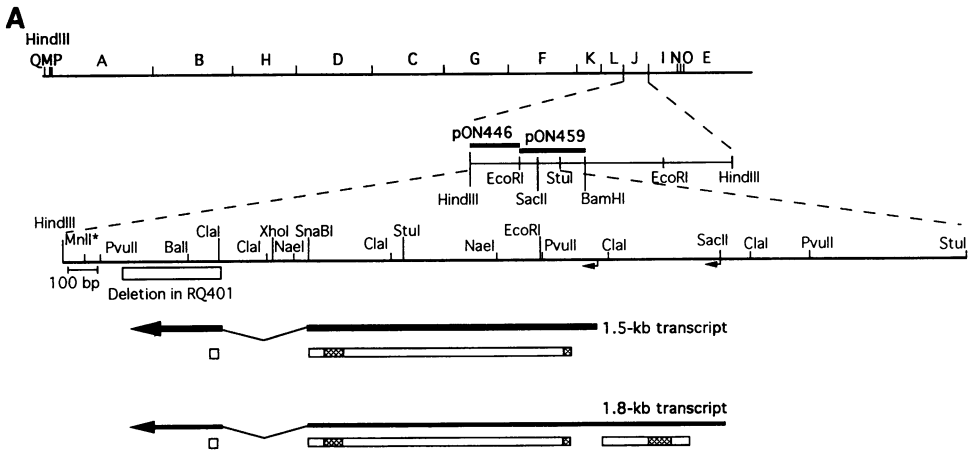
Immunoprecipitation. NIH 3T3 cells were infected with $K181^+$ or RQ401 or left uninfected (mock). Cells were incubated in methionine-free medium (Irvine Scientific) for 1 h at 7 hpi, and then [^{35}S]methionine (specific activity, 1,054 Ci/mmol; Amersham) was added at 50 μCi/ml for 2 h at 8 hpi. Cells were washed three times with PBS, resuspended in lysis buffer (10 mM Tris-Cl [pH 7.4], 1 mM EDTA, 150 mM NaCl, 0.5% [wt/vol] sodium deoxycholate, 1% [vol/vol] Nonidet P-40) containing 2 μg of aprotinin per ml (Sigma, St. Louis, Mo.) and 2 μg of leupeptin per ml (Boehringer Mannheim), and incubated on ice for 10 min. To pellet debris, the suspension was centrifuged at 18,000 $\times$ *g* in a tabletop microcentrifuge, and then the supernatant was removed to a new tube. Mouse antiserum to $K181^+$ or Smith virus was added at a 1:250 dilution and incubated with supernatant for 1 h on ice. This was followed by incubation with 30 μl of Pansorbin (Calbiochem, La Jolla, Calif.) for 1 h on ice, brief centrifugation, and four washes in lysis buffer. Immunoprecipitated proteins were separated on denaturing 12 or 15% polyacrylamide gels, and labeled proteins were detected by autoradiography.

At 31 h after transfection, COS-7 cells were incubated in methionine-free medium for 1 h, and then [^{35}S]methionine was added at 50 μCi/ml for 2 h. Cells were harvested and immunoprecipitation was performed as described above.

Generation of recombinant viruses. The construction of RM868, a recombinant virus with a *lacZ-gpt* insertion within the *sgg1* gene, is described in reference 47. In brief, at 3 hpi with murine CMV, NIH 3T3 cells were transfected by a calcium phosphate method (4) with pON868, except that incubation of the cells with calcium phosphate-DNA mixture was in a 5% CO_2 incubator. At 72 hpi, a virus stock was prepared from this culture and used to infect fresh cells. At 3 hpi, the medium in this culture was replaced with medium containing 12.5 μg of mycophenolic acid per ml and 150 μg of xanthine per ml. At 72 hpi, virus was harvested, and after three rounds of selection, recombinant virus was plaque purified as previously described (23). RM461, a recombinant virus with a *lacZ* insertion, has been described and characterized previously (44). For this virus, the *lacZ* gene under transcriptional control of the human CMV immediate-early promoter-enhancer was inserted into the *Hin*dIII site between *sgg1* and *ie2*. Prior to purification, the pool of viruses was passaged through mice, and salivary gland stocks were harvested at 14 days post-infection (dpi). The salivary gland stocks were plaque purified three times by selecting β-galactosidase-positive plaques.

Detection of RM868 in organs and salivary gland. For titration of virus growth in organs, adult male BALB/c.ByJ mice were inoculated i.p. with 10^6 PFU of tissue-culture-propagated murine CMV ($K181^+$ or RM868). Five animals were inoculated with each virus. At 3 dpi, mice were sacrificed by CO_2 asphyxiation. Organs were harvested, weighed, and sonicated as a 10% (wt/vol) tissue sonicate in Dulbecco's modified Eagle's medium-autoclaved skim milk (50:50). Twenty-five-day-old BALB/c mice were inoculated i.p. with 10^6 PFU of tissue-culture-propagated $K181^+$ or RM868. At 15 dpi, mice were sacrificed by CO_2 asphyxiation and all major salivary glands were harvested. Glands were divided in two; one half was used for titer determination, and the other half was frozen for sectioning. Plaque assays were performed with serial dilutions of a 10% suspension of each sonicated tissue. Cultures were overlayed with 0.75% carboxymethylcellulose for 5 days, and plaques were observed after fixation in absolute methanol by staining with Giemsa stain. The limit of viral detection in these assays was 10 PFU/ml.

For in situ staining, salivary glands (harvested as described above) were frozen in OCT (Miles, Elkart, Ind.) at −135°C in liquid N_2. Salivary glands were sectioned with a model 5030 microtome (Bright Instrument Company, Ltd., Huntingdon, England), and 5-μm sections were placed on glass slides and



B

```
ctgatagaccgtgaagatttacaagaactgtgacggggatcgcAATAAAatgtttttaccgacggcggcggcgtggttatttctcgcgcacggtgtcgcg
tcgcagatcaccactaccgtctcgcagacgcgaggttacactaccgcacgtaacaaaacctacatcgagtggtgcctttcgtacgaacgaaacggttcgc
                                                    *  Q  N  L  H  R  V  V  P  F  V  R  T  K  R  F  A

ttccctaaacgtaacgtggaacaggacgttgtacgacacagagatcctcttcgagacctcgtgactagattcacgacggagagatgcgtcgcacgcaag
 S  L  N  V  T  W  N  R  T  L  Y  D  T  E  I  L  F  E  T  S  W  T  R  F  T  T  E  R  C  V  A  R  K
ttggtcggagccgaggacgccatgatacagatactcagatttgatcctgacgacaaaaatatggaagagttgcgtgtgaaatttgaggggaccataaacc
L  V  G  A  E  D  A  M  I  Q  I  L  R  F  D  P  D  D  K  N  M  E  E  L  R  V  K  F  E  G  T  I  N  H
acgattcaacggggtcattcgtttcctacatcaacacacaatcaggtagatggtcggtggagaaggttatgaaggtagagatctcttttgtgtgtgaatt
  D  S  T  G  S  F  V  S  Y  I  N  T  Q  S  G  R  W  S  V  E  K  V  M  K  V  E  I  S  F  V  C  E  L
gaaagtcagtgaggatgcagccagatacagctgcgagcctgaaggctacgggtggctgaccaaaggatacttgaccgcgaccgttcgtgtgaatggaagc
 K  V  S  E  D  A  A  R  Y  S  C  E  P  E  G  Y  G  W  L  T  K  G  Y  L  T  A  T  V  R  V  N  G  S
gaacacagcatccacctaccagagccaggggcgcgtatatcgagtacgaacgggggagtttcctatcgggttacttaggacccggcgggccgctcgtca
E  H  S  I  H  L  P  E  P  G  G  A  Y  I  E  Y  E  R  G  S  F  L  S  G  Y  L  G  P  G  G  P  L  V  T
ctgtcacagtgctgcaaccgtgtccaccgagggagatcgatctgatcatcgaactgcacccgcgttacaagctgtatcagagatttgtaattccggagcg
  V  T  V  L  Q  P  C  P  P  R  E  I  D  L  I  I  E  L  H  P  R  Y  K  L  Y  Q  R  F  V  I  P  E  R
ccggggggAATAAAaacttcgcgatcctccggtgccgcggctcattgtcggacagaacagaactctgtgagagaggggacaggagaatcgcggcgaggca
 R  G  E  *                    1.8 start ┗►                                                        M   1

                                                                                                tggctattaacaccacagaatgtccctctccctctgagaccgtgtccgctcagaccatcgtgctgacctgcgccgccctgacttcggtgttgttggcggc
  2                                                                                              A  I  N  T  T  E  C  P  S  P  S  E  T  V  S  A  Q  T  I  V  L  T  C  A  A  L  T  S  V  L  L  A  A  34
                                                                                                cgctatattggtggtggcgatctccgcgctcgtgagacgctgccgacggcagcaaacccctcgcggcgacgagcagctcccgatcgaatgtagcgtctat
 35                                                                                             A  I  L  V  V  A  I  S  A  L  V  R  R  C  R  R  Q  Q  T  P  R  G  D  E  Q  L  P  I  E  C  S  V  Y  67
                                                                                                ttggtacagaacaacggggtggacgacgacgacatccccccggggtttctggggcgccccgccgggtacgaagacatgagcccgcactccttcccgctgg
 68                                                                                             L  V  Q  N  N  G  V  D  D  D  D  I  P  P  G  F  L  G  R  P  A  G  Y  E  D  M  S  P  H  S  F  P  L  D 101
                                                                                                acgagatttacgagagtatcgattgaacggctgttcgtgacgtcgactgccatccccaaaccccggtgTATAAagagtagatatcaacctcccgcatcta
102                                                                                             E  I  Y  E  S  I  D  *                                                    1.5 start ┗►  108
                                                                                                cacactttcgcacccgacaggaagaatgtcaccaccgctgctgttcacgatttctctcttttcgcgctggggtccaggccgctcgcgggctatagcaat
  1                                                                                                                M  S  P  P  L  L  F  T  I  S  L  F  F  A  L  G  S  R  P  L  A  G  Y  S  N  25
                                                                                                ttggcagtgtggtactccgattatgaaaacggaacctacagctgtgaattcgtattgccgagacccaacaacatgccgctgccgtttaccaactgctacc
 26                                                                                             L  A  V  W  Y  S  D  Y  E  N  G  T  Y  S  C  E  F  V  L  P  R  P  N  N  M  P  L  P  F  T  N  C  Y  L  59
                                                                                                ttcccgagaaaatcaaatcctcggcgacggtgacgccgaacggcacgatcttgcttcacgacaactacctgaccaccaagttggtggaggatagttcgac
 60                                                                                               P  E  K  I  K  S  S  A  T  V  T  P  N  G  T  I  L  L  H  D  N  Y  L  T  T  K  L  V  E  D  S  S  T  92
                                                                                                cggcaagcggcatcgccggctgtcgttcagagacccggagtttgagaaacgtcacccgtgtctttactgcgagagaccgggcgattcgatggatatggag
 93                                                                                             G  K  R  H  R  R  L  S  F  R  D  P  E  F  E  K  R  H  P  C  L  Y  C  E  R  P  G  D  S  M  D  M  E  125
                                                                                                agatgtctgatgtgcaaggacacgacaatggcggtcttactaaaggatttacacacgcccaaacggcgcgatttttcgcctatgttggtccgggccacgg
126                                                                                             R  C  L  M  C  K  D  T  T  M  A  V  L  L  K  D  L  H  T  P  K  R  R  D  F  S  P  M  L  V  R  A  T  V 159
                                                                                                tggagaccgtcaatcaccggccttatttcacgtgccggtttcgcgtaactccggacacccagaagatcgagacggcgtggttgaacgacgactgcacctt
160                                                                                               E  T  V  N  H  R  P  Y  F  T  C  R  F  R  V  T  P  D  T  Q  K  I  E  T  A  W  L  N  D  D  C  T  F  192
                                                                                                cggcgccaaaacgaaågatccgatcgaggcctgcgccaaaacggccagctacaagaaatcgatacgcgagggaagatacgtgaccatgagcgccaaggtg
193                                                                                             G  A  K  T  K  D  P  I  E  A  C  A  K  T  A  S  Y  K  K  S  I  R  E  G  R  Y  V  T  M  S  A  K  V  225
                                                                                                aaacccggagacccgcgcacctgccgcgtgtgcctgctgaagacggacaacaagtacgtgctgatggcgcgtccatgcctgacggacaagcggcgccagg
226                                                                                             K  P  G  D  P  R  T  C  R  V  C  L  L  K  T  D  N  K  Y  V  L  M  A  R  P  C  L  T  D  K  R  R  Q  E 259
                                                                                                aatacatccgaaacgtgcgcatcgtgcacgcgttggtgatcagcgtgtgggcggtcgcgatcgccctggtgggcttcgccgccgtgcgcctgctattaag
260                                                                                               Y  I  R  N  V  R  I  V  H  A  L  V  I  S  V  W  A  V  A  I  A  L  V  G  F  A  A  V  R  L  L  L  S  292
                                                                                                cctggacagcgtcgcccggctggtgaaaaaaatgttcgcggtacgtatactgccgttaccgttaccgctggtcgtggtgtggttcgccgccggcgacacg
293                                                                                             L  D  S  V  A  R  L  V  K  K  M  F  A                                                 307
                                                                                                gacgtgcgccgtgactgctggccgtggagagcgtcctgctctccgggctcgagctggggttcaaccatcgatccctgaacgcgtggaccgagacggggcc
                                                                                                caaatgaaactggcgctggagaggtagctcttccgagaaaacggtcgggttccgagaagggctctctctctctgactaatcaccccccggctctgacgtgt
                                                                                                gcccgtttttgtctttcaggctggatataggaggctatcttaagactatcgatgtgctggtcggctcatctgacgcacgggaggtgagacagggaagaga
306                                                                                                                A  G  Y  R  R  L  S  *                                    312
                                                                                                aaactcaggacgtcgtacgccccgtgctgggaagttacaacgcgagggtggccaagatgaccgccgacctgtccgcctcgacacaacccgggaaaaagg
                                                                                                tgttgaatgctagagtattcatcagcgtgcaggatgacgtgtgacgatgaaggtgtcgttactgcctatgggaagctggagaccccacggcctatgtata
                                                                                                tttatgtttttgcttaacgtttcactgattcacAATAAAatctcatctgaaatgggaacgctcctcgtgtgctgcctcgtctgttgcgtcgtgttggtcg
                                                                                                tgtctaccgtcgcggacctccgcgagccatgttgcgcgcggcgcagctcgcacccctctcccgctgtacgcggtaacagagcgccgaatacactaacacttt
                                                                                                cgtgcggncgagaggtggttttcactaccttctctgggga
```

FIG. 1. (A) Structure of the *sgg1* gene. The top line represents a *Hin*dIII map of the murine CMV K181$^+$ genome (27) with the *Hin*dIII J region expanded below. Plasmids, pON446 and pON459, contain genomic DNA sequence (thick lines). The *Hin*dIII-*Stu*I region is expanded with a restriction map of the *sgg1* region. Arrows below the line near the *Sac*II and *Cla*I sites indicate the start sites of transcription. The *sgg1* 1.5- and 1.8-kb spliced transcripts are shown below the restriction map (thick arrows). ORFs are represented as open boxes beneath the transcripts. Hydrophobic sequences are shown as hatched boxes within the ORFs. The deletion in RQ401 is indicated as an open box. The asterisk at *Mnl*I indicates that the site is not unique. (B) Nucleotide sequence of the *sgg1* gene. The sequence from the *Stu*I to the *Mnl*I of the murine CMV K181 *Hin*dIII J fragment is shown, along with the deduced amino acid sequence of the three ORFs (>100 aa) found in the region. ORF 1 (219 aa) is located upstream of the *sgg1* gene. The *sgg1* gene includes the second (108-aa) and third (312-aa) ORFs. Putative TATA boxes are denoted by uppercase letters, and putative polyadenylation sites are in uppercase and are underlined. A putative TATA box (AATAA) for the 1.8-kb transcript is located 29 nucleotides upstream of the start site, and a TATA box for the 1.5-kb transcript is located 19 nucleotides upstream of the start site. The transcription initiation sites for the 1.5-kb and 1.8-kb transcripts are indicated by arrows below the sequence. Two potential N-linked glycosylation sites are underlined, and three CXXC motifs are boxed. The insertion site of the *lacZ-gpt* cassette in RM868 is shown as an å in the codon of Lys-197 in the nucleotide sequence.

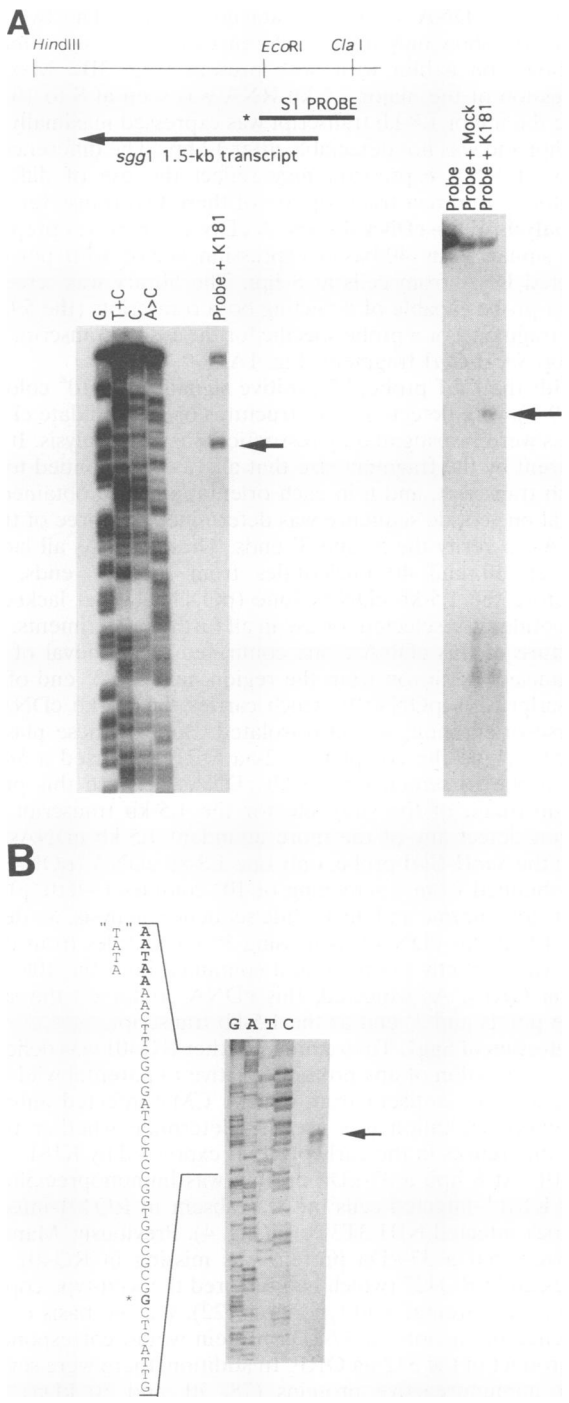


FIG. 2. Start sites of the 1.5- and 1.8-kb transcripts. (A) S1 analysis of the 5′ end of the 1.5-kb transcript. Mock RNA or RNA collected at 8 hpi with K181 virus was hybridized to the 228-bp end-labeled *Eco*RI-*Cla*I probe shown above the autoradiogram. The left set of lanes show the S1 protection analysis and Maxam-Gilbert sequencing reactions separated simultaneously on an 8% polyacrylamide–7 M urea gel. The right set of lanes show the same probe alone and hybridized with mock-infected or K181-infected RNA. The protected bands are indicated by arrows to the right of both sets of lanes. (B) Primer extension analysis of the 1.8-kb transcript. Nucleotide sequence determined by the dideoxy method is shown starting at the putative TATA box of the 1.8-kb transcript. Primer extension analysis samples were separated simultaneously to the right of the sequence ladder on an 8% polyacrylamide urea gel. With primer 970-936, a 140-bp band corresponding to the start of transcription is indicated by the arrow on the right. The band is located at the guanine CCGCGG of the *Sac*II site and is indicated by a boldface letter with an asterisk on the left.

allowed to air dry. Slides were fixed in 0.5% glutaraldehyde (Baker, Phillipsburg, N.J.) solution in PBS containing 2 mM $MgCl_2$ for 15 min. The slides were washed in PBS–2 mM $MgCl_2$ and stained with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal; 250 μg/ml; Clontech, Palo Alto, Calif.) in PBS–2 mM $MgCl_2$–100 mM $K_3Fe(CN)_6$–100 mM $K_4Fe(CN)_6$ overnight, as described previously (15). After this step, slides were stained with hematoxylin and eosin (Sigma).

Nucleotide sequence accession number. The accession number for the *sgg1* sequence is L32187.

RESULTS

Structure of the *sgg1* gene. To gain insights into the structure and organization of the *sgg1* gene and to map the 1.5- and 1.8-kb transcripts previously detected in the left portion of the *Hin*dIII J fragment of the murine CMV genome (24), we determined the nucleotide sequence of a 3,040-bp *Stu*I-to-*Mnl*I fragment constituting a portion of pON446 and pON459 from within the *Hin*dIII J fragment (Fig. 1). With the sequence, we constructed a detailed restriction map of the left end of the *Hin*dIII J fragment, identified possible transcription signals, and located three ORFs in this region (Fig. 1). All large (>100 codons) ORFs in this region are oriented in the same, leftward direction on the viral genome (Fig. 1). The coterminal 3′ ends of the major 1.5-kb and minor 1.8-kb mRNA species and the 281-nucleotide intron spliced out of both transcripts had been previously mapped (22, 24). The largest ORF within the transcribed region is 312 aa in size. We used a hydropathy plot (9) to predict a signal sequence at the amino terminus (aa 1 to 17) and a transmembrane domain at aa 268 to 287 (Fig. 1). Two putative asparagine (N)-linked glycosylation sites were present at aa 35 and aa 72 (Fig. 1B). Preliminary data suggested that at least one site is glycosylated, because treatment of COS-7 cells transfected with pON4014 (the 1.5-kb cDNA) with *N*-glycosidase F results in a reduction in the size of Sgg1 from 37 kDa to 34 kDa (data not shown). The predicted ORF is arginine rich (9%) and contains three CXXC motifs. These characteristics suggest that *sgg1* most likely encodes a type 1 integral membrane protein and that the protein may contain a metal binding region.

The start site of the 1.5-kb transcript was determined by S1 nuclease protection with a 228-bp end-labeled *Eco*RI-*Cla*I fragment. A single band of 160 nucleotides was protected, indicating that the 5′ end of the 1.5-kb transcript was 160 nucleotides upstream of a unique *Eco*RI site in the region (Fig. 2A). Mock-infected cellular RNA did not give a signal in this region (Fig. 2A and reference 22). Primer extension was used to locate the start site for the 1.8-kb transcript. A single band corresponding to the start of the 1.8-kb transcript was located at the last guanine, CCGCGG (represented by boldface type), of the *Sac*II site (Fig. 2B). Thus, the 1.8-kb message is expressed from a promoter positioned approximately 450 bp upstream of the promoter driving expression of the 1.5-kb mRNA. The 1.8-kb cDNA recovered from the library is within 30 nucleotides of this predicted start site for the 1.8-kb transcript. On the basis of the position of the start site, the 1.8-kb transcript appears to use an AATAAA sequence as a TATA element. This signal could also be a polyadenylation signal for an upstream gene (47).

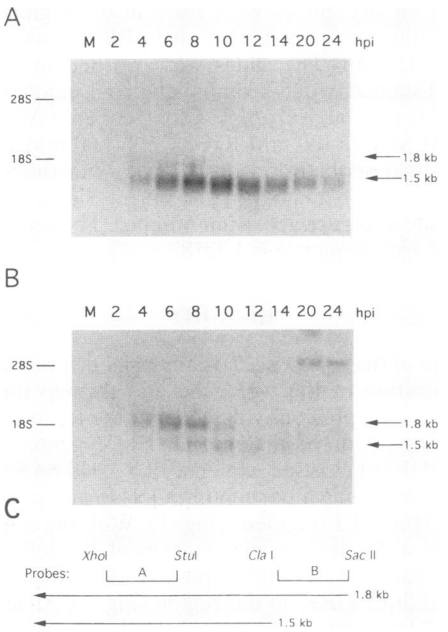


FIG. 3. RNA blot hybridization analysis of the kinetics of expression of *sgg1*. RNA was harvested from NIH 3T3 cells at various times between 2 and 24 hpi. Whole-cell RNA was separated, transferred to GeneScreen Plus, and hybridized sequentially with two *sgg1*-specific ^{32}P-labeled probes. (A) RNA was probed with a 435-bp *Stu*I-*Xho*I probe that spans a region common to both the 1.5- and 1.8-kb transcripts. 28S rRNA and 18S rRNA are indicated on the left, and the 1.8- and 1.5-kb transcripts are indicated on the right by arrows. (B) The blot used in panel A was hybridized with a 383-bp *Sac*II-*Cla*I probe that was specific for the 1.8-kb transcript. The 1.5-kb transcript is still visible (although light) after this hybridization because the first probe is still present. At late time points 20 and 24 hpi, a 4.7-kb band is observed and may represent background hybridization in the region of large rRNA. (C) Diagram of probes used in panels A and B. Probe A, an *Xho*I-*Stu*I fragment, spans a region common to both the 1.5- and 1.8-kb transcripts. Probe B, a *Sac*II-*Cla*I fragment is upstream of the 1.5-kb start site.

The 1.8-kb transcript contains an additional 108-aa ORF located in the region upstream of the 312-aa ORF. Stop codons are found in all three reading frames in the 98 bp between the two ORFs. The predicted 108-aa ORF also has a hydrophobic region, between aa 25 and 45, that is potentially capable of spanning a membrane (Fig. 1).

The deletion in RQ401 (Fig. 1A) eliminates the putative polyadenylation signal shared by both transcripts and probably results in instability of the 1.5- and 1.8-kb mRNAs (24). A large predicted ORF is positioned upstream of the *sgg1* region (called HJ4 in reference 47), as depicted in Fig. 1B; however, transcripts have not been detected in this region (47).

***sgg1* RNA accumulation.** We examined steady-state RNA levels of the 1.5- and 1.8-kb transcripts at various times postinfection. RNA from K181$^+$-infected NIH 3T3 cells harvested between 2 and 24 hpi was quantitated with two *sgg1*-specific probes, a labeled 435-bp *Stu*I-*Xho*I fragment that detects both transcripts and a 383-bp *Sac*II-*Cla*I fragment that specifically detects the 1.8-kb transcript (Fig. 3 and reference 22). Both RNAs were expressed at early times (4 hpi), and the 1.5-kb RNA continued to be expressed through late times (24 hpi), although its abundance was reduced. Consistent with both transcripts being in the β class, their expression was insensitive to the addition of phosphonoformic acid, an inhibitor of viral DNA synthesis (data not shown). The two *sgg1* transcripts show only minor differences in expression kinetics as shown on a blot with both present (Fig. 3B). Maximal expression of the major 1.5-kb RNA was seen at 8 to 10 hpi, while the minor 1.8-kb transcript was expressed maximally at 6 to 8 hpi and was not detectable after 12 hpi. The differences in levels of RNA expression may reflect the use of different promoters to direct transcription of these two transcripts.

Analysis of the cDNA library. A cDNA library was prepared in a simian virus 40-based expression vector with poly(A)-selected RNA from cells at 8 hpi. The library was screened with a probe capable of detecting both transcripts (the 592-bp *Cla*I fragment) or a probe specific for the 1.8-kb transcript (the 383-bp *Sac*II-*Cla*I fragment; Fig. 1A).

With the *Cla*I probe, 87 positive signals from 10^4 colonies (~0.9%) were detected. The structures of 12 candidate cDNA clones were investigated by restriction enzyme analysis. It was apparent by the fragment size that all 12 corresponded to the 1.5-kb transcript, and 6 in each orientation were obtained. A partial nucleotide sequence was determined for three of these cDNAs to verify the 5′ and 3′ ends. These cDNAs all lacked between 30 and 40 nucleotides from their 5′ ends, and therefore the 1.5-kb cDNA clone (pON4014) that lacked 30 nucleotides was selected for use in all further experiments. The structure of this cDNA clone confirmed the removal of one 281-nucleotide intron from the region near the 3′ end of the transcript (24). pON4019, which carries the 1.5-kb cDNA in reverse orientation, was also isolated. Both of these plasmid clones include the complete 312-aa ORF. We used a *Sac*II-*Cla*I probe to screen for 1.8-kb cDNAs because this probe was upstream of the start site for the 1.5-kb transcript and did not detect any of the more abundant 1.5-kb cDNAs. By using the *Sac*II-*Cla*I probe, only one 1.8-kb cDNA (pON4015) was obtained from a screening of 10^4 colonies (~0.01%). By restriction enzyme and nucleotide sequence analysis, we determined that this cDNA was missing 30 nucleotides from its 5′ end, was correctly oriented, and contained both the 108- and 312-aa ORFs. As expected, this cDNA contained the same splice points and 3′ end as the 1.5-kb transcript.

Detection of Sgg1. To examine whether RQ401 was deficient in the expression of any proteins relative to parental wild-type virus, we used antisera from murine CMV-infected animals. Immunoprecipitation was used to determine whether there were differences in the early proteins expressed by K181$^+$ and RQ401. At 8 hpi, a 37-kDa protein was immunoprecipitated from K181$^+$-infected cells but was absent in RQ401-infected or mock-infected NIH 3T3 cells (Fig. 4). Previously, Manning reported that a 37-kDa protein was missing in RQ401 and expressed in RQ427 (which had restored the wild-type copy of *sgg1*) as in parental wild-type virus (22). On the basis of our sequence predictions, a 37-kDa protein would correspond to the product of the 312-aa ORF. In addition, there were several other immunoreactive proteins (78, 70, and 50 kDa) that appeared to be expressed at reduced levels in the *sgg1* mutant viruses RQ401 (22) and RM868 (data not shown). The larger proteins may be the products of other genes that are regulated by or expressed in association with *sgg1*. Consistent with the rescue of the growth properties, all of these proteins were present at the same levels in RQ427. Thus, the 37-kDa protein is apparently encoded by the *sgg1* gene and is absent in *sgg1* mutants.

In order to confirm that the 37-kDa protein was the product of the *sgg1* gene, we introduced the expression clone pON4014 or pON4015 (1.5- and 1.8-kb cDNAs) into COS-7 cells and employed immunoprecipitation analysis with the mouse anti-K181$^+$ antiserum. Both constructs expressed abundant levels

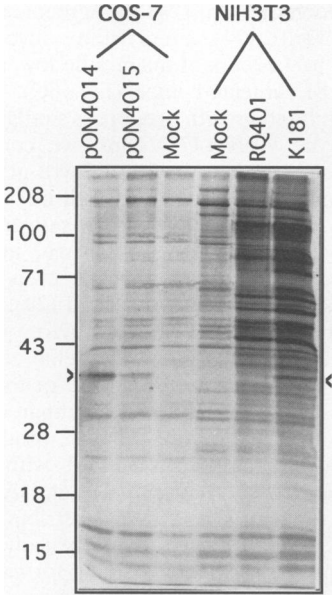


FIG. 4. Immunoprecipitation of the 37-kDa *sgg1* protein with mouse antisera raised by infection with murine CMV. The left three lanes show immunoprecipitated proteins from COS-7 cells that were transiently transfected with expression plasmids, pON4014 or pON4015 or that were mock transfected and then labeled with [^{35}S]methionine between 32 and 34 h posttransfection. The right three lanes show proteins immunoprecipitated from NIH 3T3 cells that were infected with K181$^+$ or RQ401, or mock infected, and labeled from 8 to 10 hpi with [^{35}S]methionine. Proteins were separated by SDS-PAGE in a 15% polyacrylamide gel. Molecular weight standards (in thousands) appear on the left. The position of the 37-kDa protein is indicated (arrowheads).

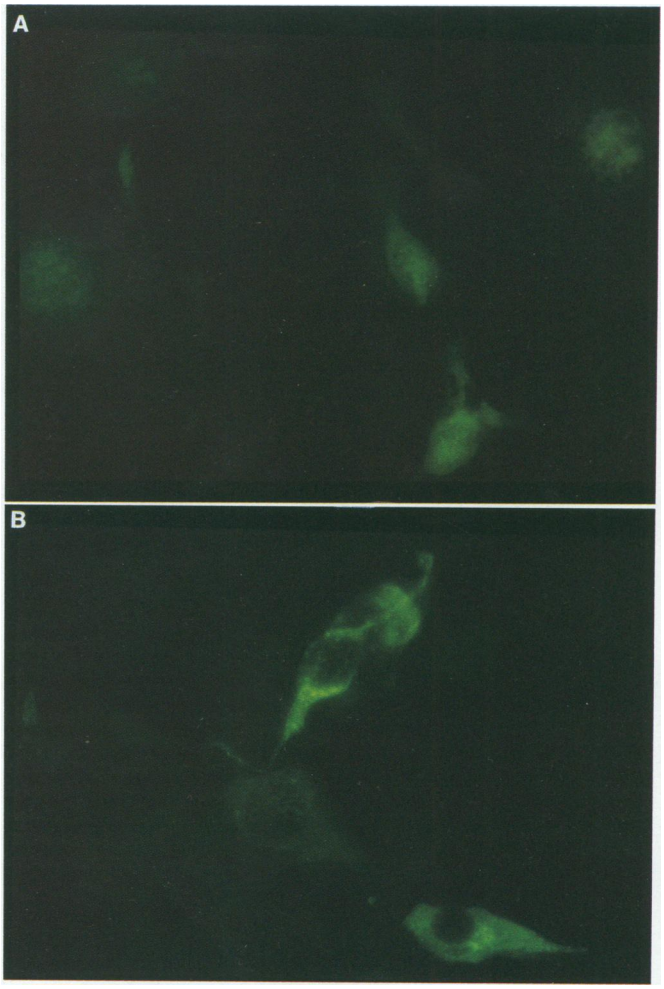


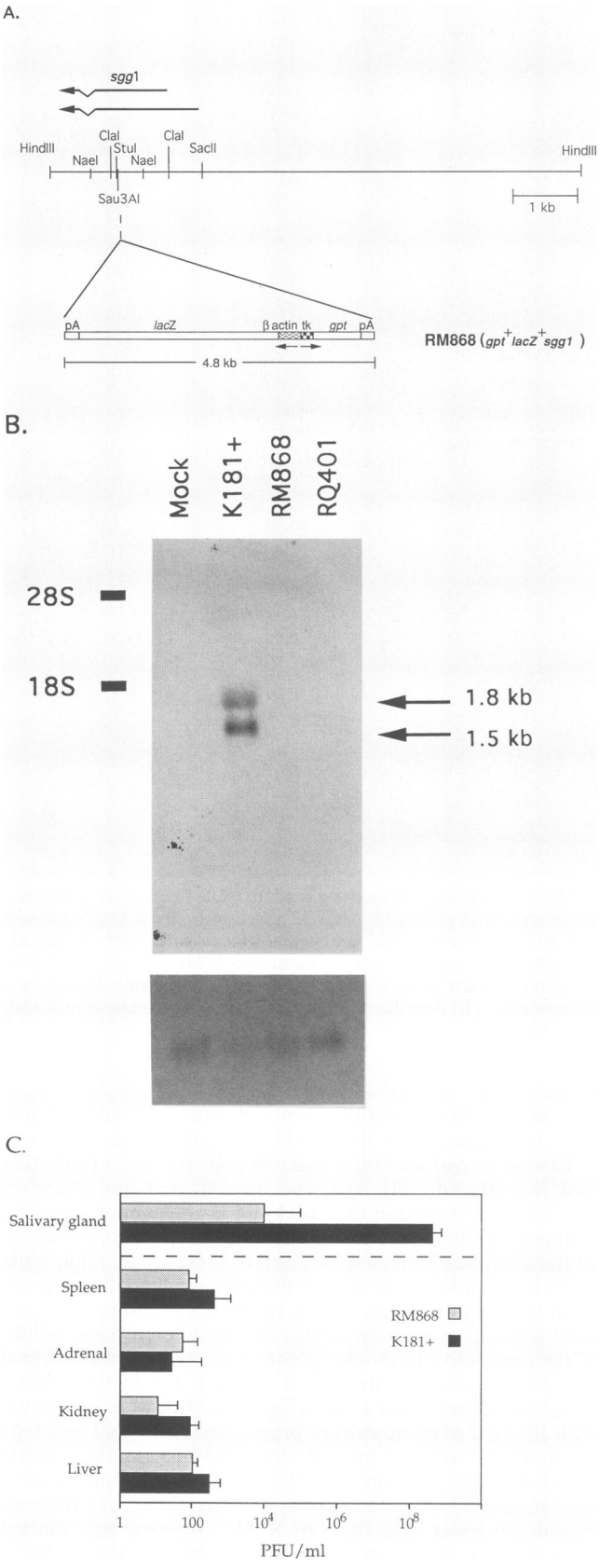
FIG. 5. Detection of Sgg1 expression. COS-7 cells were transiently transfected with pON4014. After 32 h, cells were fixed, incubated with anti-RQ401 or anti-K181$^+$ antiserum, and revealed with fluorescein isothiocyanate goat anti-mouse immunoglobulin G. (A) Anti-RQ401 antiserum, which lacks anti-*sgg1* reactivity, showed only background fluorescence. (B) Anti-K181$^+$ antiserum showed punctate cytoplasmic fluorescence.

of RNA as detected by RNA blot analysis (data not shown), and both expressed an immunoreactive 37-kDa protein detected by immunoprecipitation with this antiserum, although expression was much more abundant when the 1.5-kb cDNA clone was introduced than when the 1.8-kb cDNA clone was used (Fig. 4). The presence of the 108-aa ORF upstream of the 312-aa ORF did not completely block expression of the 37-kDa protein, although we do not know whether this short ORF is translated from the 1.8-kb transcript. We examined immunoprecipitates by SDS-polyacrylamide gel electrophoresis (PAGE) on 20% gels but did not detect any proteins of the size (12 kDa) predicted to correspond to the product of the 108-aa ORF (data not shown).

Subcellular localization of Sgg1. The 312-aa Sgg1 contains a short hydrophobic signal sequence, two N-linked glycosylation sites, and a transmembrane domain, suggesting it might be a membrane glycoprotein. We used COS-7 cells to examine the subcellular localization of Sgg1. After transfection with pON4014 (the 1.5-kb cDNA), punctate-to-diffuse cytoplasmic fluorescence was observed in the COS-7 cells (Fig. 5). Thus, the cytoplasmic localization of the protein was consistent with the sequence characteristics of Sgg1.

Recombinant virus structure and gene expression. RM868 carries a *lacZ-gpt* cassette disrupting the 312-aa ORF (47). The insertion was mapped by restriction enzyme analysis to a *Nae*I fragment (Fig. 6A). We determined the nucleotide sequence of a portion of pON868 to define the site of insertion of the cassette into *sgg1*. The *lacZ-gpt* cassette is inserted at a *Sau*3AI site following Lys-197 of the *sgg1* ORF (Fig. 1B) and disrupts Sgg1 after Pro-199. This insertion interrupts expression of the 1.5- and 1.8-kb transcripts (Fig. 6B and reference 47) and prevents expression of the 37-kDa protein (data not shown).

Growth of *sgg1* mutant viruses at primary sites of infection and in cultured cells. We have investigated the course of murine CMV infection in 3-week-old BALB/c mice following intraperitoneal inoculation of virus. Peak replication in tissues of the abdomen such as spleen and liver occurs at 3 dpi, while titers in secondary sites such as the salivary gland do not reach peak levels until 14 dpi, after a viremia. Therefore, we examined replication in these tissues at two different time points. Virus replication in mouse spleen, liver, kidney, and adrenal gland tissue was examined at 3 dpi and that in the salivary gland was examined at 15 dpi (Fig. 6C). Mice were inoculated with 10^6 PFU of tissue-culture-propagated virus from wild-type strain K181$^+$ or *sgg1* mutant strain RM868. In the adrenal gland, RM868 replicated as well as K181$^+$, while titers in liver, spleen, and kidney tissue were approximately 0.5 to 1 order of magnitude lower than titers of K181$^+$. Another *sgg1* mutant virus, RQ401, had titers comparable to those of the wild-type virus in the liver and exhibited approximately 10-fold lower



titers in the spleen at 3 dpi (24). In agreement with previous results from RQ401 (24), titers from salivary glands from RM868 were 3 to 4 orders of magnitude lower at 15 dpi than those from K181$^+$-infected mice (Fig. 6C). In cell culture, RM868 also exhibited growth properties similar to those of the parental wild-type virus (47). Thus, we confirm that *sgg1* appears to be dispensable for viral growth in most organs of the infected mouse as well as in cell culture.

Localization of *sgg1* mutant virus in salivary glands. Manning et al. previously reported that an *sgg1* mutant virus was capable of dissemination via leukocytes even though replication was depressed in the salivary gland (24). To address the question of whether an *sgg1* mutant virus was capable of infecting the same salivary gland epithelial cell type as wild-type virus and to investigate the nature of the block to viral replication in salivary gland, we compared RM868 and RM461, which is the wild type for *sgg1*. Stoddart et al. (44) reported that recombinant virus RM461, with an insertion of *lacZ* into the *Hin*dIII site between *sgg1* and *ie2*, had the same cellular site of replication and tissue distribution as the wild-type virus. RM868 can be detected in the salivary gland in the serous acinar epithelial cells by staining for *lacZ* activity (Fig. 7A), as can RM461, which carried a wild-type *sgg1* gene (Fig. 7B). Thus, *sgg1*-deficient viruses maintain the ability to disseminate and enter the appropriate target cells in the salivary gland even though they lack the ability to replicate efficiently once there.

DISCUSSION

We have characterized a murine CMV β gene, *sgg1*, that is dispensable for growth in cell culture but dramatically impacts virus growth in salivary glands. Deletion or insertion mutations in the *sgg1* gene result in viruses that replicate very poorly in the salivary gland. In contrast to the specific defect in *sgg1* mutants, other viral mutants that exhibit decreased growth in the salivary gland are also very deficient for growth in other organs such as the liver, spleen, and kidney and peripheral blood leukocytes (2, 46). Thus, the *sgg1* defect appears to be specific to virus growth in acinar epithelial cells in the salivary glands and does not appear to be the result of a general replication defect.

FIG. 6. Recombinant virus structure, transcription, and phenotype. (A) Insertion site of RM868 in the *Hin*dIII J fragment. The *Hin*dIII J fragment is shown with the two *sgg1* transcripts just above on the left. *Cla*I, *Sac*II, *Nae*I, and *Stu*I restriction enzyme sites are noted in the *sgg1* region (for more detail of the region, see Fig. 1A). The insertion of the *lacZ-gpt* cassette (47) is located at the *Sau*3AI site adjacent to the *Stu*I site in *sgg1* (also noted in Fig. 1B). The cassette contains *lacZ* under control of the β-actin promoter and *gpt* under control of the thymidine kinase (tk) promoter. Simian virus 40 poly(A) (pA) sites on the cassette are noted. The phenotype of recombinant virus RM868 is shown to the right of the insertion. (B) RNA blot hybridization analysis of *sgg1* transcripts. Whole-cell RNA extracted from NIH 3T3 cells infected 8 h previously with K181$^+$, RM868, or RQ401 virus or mock-infected was separated, transferred to GeneScreen Plus, and hybridized with an *sgg1*-specific ^{32}P-labeled *Eco*RI-*Hin*dIII probe. The probe spans a region common to both transcripts. The blot was stripped and then hybridized with a ^{32}P-labeled β-actin probe as an internal RNA control. The 1.9-kb β-actin band detectable in all lanes is shown below the gel. (C) Characterization of RM868 versus K181$^+$ virus in various tissues. BALB/c mice were infected with 10^6 PFU of tissue-culture-propagated RM868 or K181$^+$ virus. The salivary glands were harvested at 15 dpi. The spleen, adrenal gland, kidney, and liver were harvested at 3 dpi. Titers are expressed as the geometric mean ± the standard deviation of the mean of the titer in a 10% tissue sonicate.

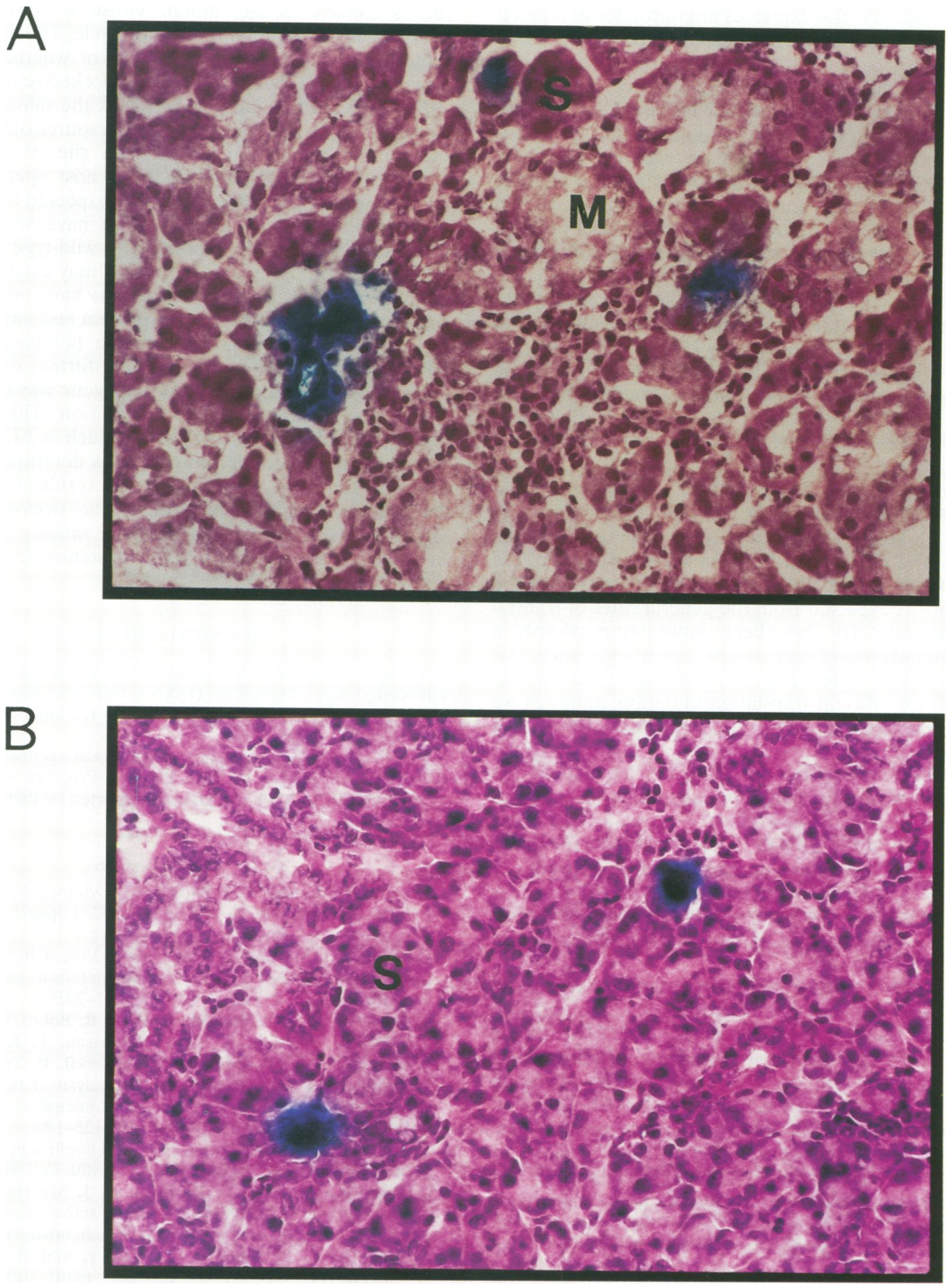


FIG. 7. Salivary gland serous acinar cells infected with RM868 or RM461 recombinant viruses. All salivary glands were harvested and frozen in liquid N_2. Five-micrometer sections were fixed and stained with X-Gal, followed by hematoxylin and eosin. (A) Two distinct cell types are visible and indicated with labels: the mucous (M) acinar cells, which are larger and vacuolar, and the serous (S) acinar cells, which are smaller and contain a round darkly staining nucleus. RM868 was seen only in serous acinar cells, as evidenced by β-galactosidase activity (shown at a ×40 magnification) staining an acinus and two individual acinar cells. (B) As a positive control, RM461, a *lacZ*$^+$ *sgg1*$^+$ virus is shown at ×40 magnification. Only serous (S) acinar cells are visible in this section. Virus is present in two serous acinar cells, as evidenced by β-galactosidase activity.

β proteins of herpesviruses typically have been associated with viral DNA replication and regulation of gene expression. The *sgg1* gene is expressed as a β gene, suggesting that it functions at early rather than late times of infection. Sgg1 has characteristics of a type 1 membrane glycoprotein, and preliminary analysis has suggested that Sgg1 is a nonstructural protein. Although glycoproteins expressed at late times in herpesvirus infections include those targeted to the viral

envelope, human CMV also expresses glycoproteins as α and β gene products (7, 43) that do not appear to be incorporated into the envelope.

Comparisons of the Sgg1 protein sequence to sequences in GenBank have not revealed any significant homology to other viral or cellular proteins. Murine and human CMVs have positional homologs that do not share significant primary amino acid sequence identity (28, 30, 48). Thus, the putative glycoproteins that are adjacent to the *ie1/ie2* locus of human CMV, such as IRL11, IRL12, or the large number of putative glycoproteins in the unique short component (3) remain possible positional homologs.

Other herpesvirus genes, such as herpes simplex virus type 1 ICP34.5, have been identified as playing a cell-specific role in tropism and virulence (6). Perhaps *sgg1* is unique to murine CMV and provides some specialized function for growth in acinar cells of the salivary gland. The localization of human CMV to ductal cells of the salivary glands suggests that these viruses differ with respect to tropism for particular subsets of epithelial cells.

By analyzing cDNAs, we found that both *sgg1* transcripts had the capacity to express the 37-kDa protein, although expression from the 1.8-kb cDNA was much lower than that from the 1.5-kb cDNA. The presence of a 108-aa ORF upstream of the 312-aa ORF in the 1.8-kb RNA decreased expression of the 37-kDa protein, as expected from the general rules of translation (21). Upstream ORFs have been implicated previously in the temporal regulation of human CMV gene expression (12, 13, 38). The 108-aa ORF has an excellent context AUG (GGCAUGG) for efficient translation initiation, although we have not yet detected expression of this protein. Although it has methionines and should be radiolabeled, this polypeptide may not be sufficiently immunogenic to be detected by the antiserum we employed for immunoprecipitation. It is important to note that the phenotype of RM868 argues against a role for the short ORF in the growth defect. In this recombinant virus, the insertion mutation was introduced downstream of the 108-aa ORF and thus would be predicted to leave its expression intact. Thus, we believe that the putative 108-aa protein does not play a role in the growth deficiency we have observed with *sgg1* mutants.

Besides Sgg1, other early proteins were expressed at reduced levels in cells that have been infected with an *sgg1* mutant. Two of these proteins, one with an apparent size of 70 kDa and the other with a size of 50 kDa, were consistently observed in different *sgg1* mutants and when different batches of mouse antiserum were employed (not shown). We do not believe that either of these represents additional *sgg1* gene products because their expression was not observed in COS-7 cells transfected with the 1.5- or 1.8-kb cDNA. This result suggests that these are viral proteins whose expression is influenced by Sgg1. Thus, Sgg1 may be a regulatory protein that influences the synthesis or stability of proteins expressed from other viral genes. In human CMV, at least two α proteins, US3 and UL37-38, encode putative glycoproteins with demonstrated regulatory properties (7). Alternatively, Sgg1 may form a multimeric complex with these proteins; thus, the loss of expression of *sgg1* may reduce the stability of these proteins.

Manning et al. (24) reported that the sgg1 mutation did not reduce the level of leukocyte-associated viremia that is responsible for viral dissemination during acute infection. Here, we extended the analysis of viral dissemination and showed that an *sgg1* mutant retains the ability to enter serous acinar cells in the salivary gland. The *lacZ* reporter gene carried by RM868 has been shown to be expressed as an α gene in mouse fibroblasts (47). Thus, our results suggest that after entry into serous acinar cells, *sgg1* mutant viruses can progress through the earliest stages of the replication cycle. The degree to which replication is reduced in these cells or whether a particular stage of the viral replication cycle is blocked remains unknown.

For both murine and human CMV, the salivary gland plays a central role in the epidemiology as a source of virus spread in the normal population. Saliva is a site of persistent and reactivated viral shedding in both host species. The *sgg1* mutants we have described would be expected to have reduced titers of virus in the saliva and to have severely reduced transmission capacity compared with wild-type murine CMV. While a functional homolog of Sgg1 may exist in the human CMV genome, sequence comparisons have not revealed any significant identity between this protein and any human CMV (AD169) ORF. It is possible that a homolog to Sgg1 will emerge when functional domains are further defined through mutagenesis. The murine CMV *sgg1* gene was first discovered as the result of a spontaneous mutation. This suggests that highly passaged human CMV strains such as AD169, the strain in which the nucleotide sequence was determined, may carry deletions that have accumulated as a result of propagation in cell culture. Understanding the tissue tropism of CMV in important target organs such as the salivary gland provides insights into the basis of CMV persistence, transmission, and disease.

ACKNOWLEDGMENTS

We appreciate the technical support of Maria Kirichenko (Stanford University) and Jeanne Luh (DNAX Research Institute). We thank Mario Roederer and Richard Greaves for reading the manuscript and for thoughtful suggestions.

L.A.L. is a predoctoral trainee and J.V. was a postdoctoral trainee in the Host Parasite Interactions Training Program supported by the PHS (T32AI07328). This work was supported by PHS grants AI30363 and AI28341.

REFERENCES

1. **Alford, C. A., and W. J. Britt.** 1993. Cytomegalovirus, p. 227–255. *In* B. Roizman, R. J. Whitley, and C. Lopez (ed.), The human herpesviruses. Raven Press, New York.
2. **Boname, J. M., and J. K. Chantler.** 1992. Characterization of a strain of murine cytomegalovirus which fails to grow in the salivary glands of mice. J. Gen. Virol. **73:**2021–2029.
3. **Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. I. Hutchison, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. M. Weston, and B. G. Barrell.** 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr. Top. Microbiol. Immunol. **154:**125–170.
4. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7:**2745–2752.
5. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5299.
6. **Chou, J., E. R. Kern, R. J. Whitley, and B. Roizman.** 1990. Mapping of herpes simplex virus-1 neurovirulence to gamma 134.5, a gene nonessential for growth in culture. Science **250:**1262–1266.
7. **Colberg-Poley, A. M., L. D. Santomenna, P. P. Harlow, P. A. Benfield, and D. J. Tenney.** 1992. Human cytomegalovirus US3 and UL36-38 immediate-early proteins regulate gene expression. J. Virol. **66:**95–105.
8. **Cowdry, E. V., and G. H. Scott.** 1935. Nuclear inclusions suggestive of virus action in salivary glands of the monkey *Cebus fatuellus*. Proc. Soc. Exp. Biol. Med. **32:**709–711.
9. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15:**321–353.

10. **Frith, C. H., and J. W. Townsend.** 1985. Histology and ultrastructure, salivary glands, mouse, p. 177–184. *In* T. C. Jones, U. Mohr, and R. D. Hunt (ed.), Digestive system: monographs on pathology of laboratory animals. Springer-Verlag, New York.
11. **Geballe, A. P., F. S. Leach, and E. S. Mocarski.** 1986. Regulation of cytomegalovirus late gene expression: γ genes are controlled by posttranscriptional events. J. Virol. **57:**864–874.
12. **Geballe, A. P., and E. S. Mocarski.** 1988. Translational control of cytomegalovirus gene expression is mediated by upstream AUG codons. J. Virol. **62:**3334–3340.
13. **Geballe, A. P., R. R. Spaete, and E. S. Mocarski.** 1986. A cis-acting element within the 5′ leader of a cytomegalovirus β transcript determines kinetic class. Cell **46:**865–872.
14. **Henson, D., and A. J. Strano.** 1972. Mouse cytomegalovirus. Clin. J. Pathol. **68:**183–195.
15. **Ho, D. Y., and E. S. Mocarski.** 1988. β-Galactosidase as a marker in the peripheral and neural tissues of the herpes simplex virus-infected mouse. Virology **167:**279–283.
16. **Ho, D. Y., and E. S. Mocarski.** 1989. Herpes simplex virus latent RNA (LAT) is not required for latent infection in the mouse. Proc. Natl. Acad. Sci. USA **86:**7596–7600.
17. **Ho, M.** 1991. Cytomegalovirus: biology and infection, 2nd ed. Plenum Publishing Corp., New York.
18. **Hudson, J. B.** 1979. The murine cytomegalovirus as a model for the study of viral pathogenesis and persistent infections. Arch. Virol. **62:**1–29.
19. **Jordan, M. C., and J. L. Takagi.** 1983. Virulence characteristics of murine cytomegalovirus in cell and organ cultures. Infect. Immun. **41:**841–843.
20. **Koszinowski, U. H., M. Del Val, and M. J. Reddehase.** 1990. Cellular and molecular basis of the protective immune response to cytomegalovirus infection. Curr. Top. Microbiol. Immunol. **154:** 189–220.
21. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
22. **Manning, W. C.** 1990. Ph.D. thesis. Stanford University, Stanford, Calif.
23. **Manning, W. C., and E. S. Mocarski.** 1988. Insertional mutagenesis of the murine cytomegalovirus genome: one prominent alpha gene (ie2) is dispensable for growth. Virology **167:**477–484.
24. **Manning, W. C., C. A. Stoddart, L. A. Lagenaur, G. B. Abenes, and E. S. Mocarski.** 1992. Cytomegalovirus determinant of replication in salivary glands. J. Virol. **66:**3794–3802.
25. **Maruyama, K.** 1993. Unpublished observation.
26. **McCordock, H. A., and M. G. Smith.** 1936. The visceral lesions produced in mice by salivary gland virus of mice. J. Exp. Med. **63:**303–310.
27. **Mercer, J. A., J. R. Marks, and D. H. Spector.** 1983. Molecular cloning and restriction endonuclease mapping of the murine cytomegalovirus genome (Smith strain). Virology **129:**94–106.
28. **Messerle, M., B. Bühler, G. M. Keil, and U. H. Koszinowski.** 1992. Structural organization, expression, and functional characterization of the murine cytomegalovirus immediate-early gene 3. J. Virol. **66:**27–36.
29. **Messerle, M., G. M. Keil, and U. H. Koszinowski.** 1991. Structure and expression of murine cytomegalovirus immediate-early gene 2. J. Virol. **65:**1638–1643.
30. **Messerle, M., G. M. Keil, K. Schneider, and U. H. Koszinowski.** 1992. Characterization of the murine cytomegalovirus genes encoding the major DNA binding protein and the ICP18.5 homolog. Virology **191:**355–367.
31. **Mims, C. A., and J. Gould.** 1979. Infection of salivary glands, kidneys, adrenals, ovaries and epithelia by murine cytomegalovirus. J. Med. Microbiol. **12:**113–122.
32. **Mocarski, E. S.** 1993. Cytomegalovirus biology and replication, p. 173–226. *In* B. Roizman, R. Whitley, and C. Lopez (ed.), The human herpesviruses. Raven Press, New York.
33. **Osborn, J. E.** 1982. Cytomegaloviruses and other herpesviruses, p. 267–293. *In* H. L. Foster, J. D. Small, and J. G. Fox (ed.), The mouse in biomedical research. Academic Press, New York.
34. **Osborn, J. E., and D. L. Walker.** 1971. Virulence and attenuation of murine cytomegalovirus. Infect. Immun. **3:**228–236.
35. **Plummer, G.** 1973. Cytomegaloviruses of man and animals. Prog. Med. Virol. **15:**92–125.
35a. **Promega.** 1989. Promega protocols and applications guide. Promega, Madison, Wis.
36. **Rowe, W. P., J. W. Hartley, H. G. Cramblett, and F. M. Mastrota.** 1958. Detection of human salivary gland virus in the mouth and urine of children. Am. J. Hyg. **67:**57–65.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Schleiss, M. R., C. R. Degnin, and A. P. Geballe.** 1991. Translational control of human cytomegalovirus gp48 expression. J. Virol. **65:**6782–6789.
39. **Smith, J. F.** 1966. The development and the histology of the salivary glands, p. 1–18. *In* Histology of salivary gland lesions. J. B. Lippincott, Philadelphia.
40. **Smith, M. G.** 1954. Propagation of salivary gland virus of the mouse in tissue cultures. Proc. Soc. Exp. Biol. Med. **86:**435–440.
41. **Smith, M. G.** 1956. Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SGV) disease. Proc. Soc. Exp. Biol. Med. **92:**424–430.
42. **Spaete, R. R., and E. S. Mocarski.** 1985. Regulation of cytomegalovirus gene expression: α and β promoters are *trans* activated by viral functions in permissive human fibroblasts. J. Virol. **56:**135–143.
43. **Stinski, M. F.** 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. J. Virol. **23:**751–767.
44. **Stoddart, C. A., R. D. Cardin, J. M. Boname, W. C. Manning, G. B. Abenes, and E. S. Mocarski.** 1994. Peripheral blood mononuclear phagocytes mediate dissemination of murine cytomegalovirus. J. Virol. **68:**6243–6253.
45. **Takebe, Y., M. Seiki, J. Fujisawa, P. Hoy, K. Yokota, K. Arai, M. Yoshida, and N. Arai.** 1988. SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. Mol. Cell. Biol. **8:**466–472.
46. **Tonari, Y., and Y. Minamishima.** 1983. Pathogenicity and immunogenicity of temperature-sensitive mutants of murine cytomegalovirus. J. Gen. Virol. **64:**1983–1990.
47. **Vieira, J., H. E. Farrell, W. D. Rawlinson, and E. S. Mocarski.** 1994. Genes in the *Hin*dIII J fragment of the murine cytomegalovirus genome are dispensable for growth in cultured cells: insertion mutagenesis with a *lacZ/gpt* cassette. J. Virol. **68:**4837–4846.
48. **Xu, J., P. B. Dallas, P. A. Lyons, G. R. Shellam, and A. A. Scalzo.** 1992. Identification of the glycoprotein H gene of murine cytomegalovirus. J. Gen. Virol. **73:**1849–1854.